(19)日本国特許庁（ＪＰ）　(12) **公開特許公報（Ａ）**　(11)特許出願公開番号

**特開平11－154058**

(43)公開日　平成11年(1999)６月８日

| (51)Int.Cl.[6] | 識別記号 | ＦＩ | |
|---|---|---|---|
| Ｇ０６Ｆ　3/06 | ３０５ | Ｇ０６Ｆ　3/06 | ３０５Ｃ |
| | ５４０ | | ５４０ |

審査請求　未請求　請求項の数５　ＯＬ　（全　９　頁）

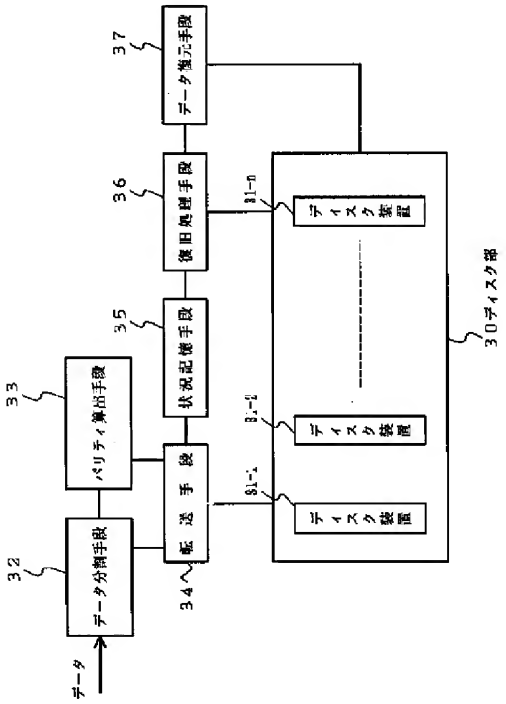
(21)出願番号　特願平9－319359

(22)出願日　平成９年(1997)11月20日

(71)出願人　000003078
株式会社東芝
神奈川県川崎市幸区堀川町72番地
(72)発明者　中村　寛男
東京都府中市東芝町１番地　株式会社東芝府中工場内
(72)発明者　笹本　享一
東京都府中市東芝町１番地　株式会社東芝府中工場内
(74)代理人　弁理士　本田　崇

(54)【発明の名称】　ディスクアレイ装置及びデータ保守方法

(57)【要約】
【課題】　データ書込途中に電源断が生じても、データを適切に復元する。
【解決手段】　データを記憶する複数のディスク装置３１－１～３１－ｎと、記憶すべきデータを所定の長さに分割するデータ分割手段３２と、分割した所定数のデータにより構成されるパリティグループについてパリティデータを算出するパリティ算出手段３３と、前記分割データ及びパリティデータを前記複数のディスク装置３１－１～３１－ｎへ転送する転送手段３４と、前記記憶すべきデータの格納先アドレス、サイズ、転送状況に関する情報を検出し記憶する状況記憶手段３５と、復旧の際に前記状況記憶手段３５の情報を検索して転送が未完を検出し、当該格納先アドレス及びサイズに基づき転送が未完となっている領域に係るパリティグループについて再度パリティデータを作成して、該当するディスク装置に書き込む復旧処理手段３６とを具備する。



【特許請求の範囲】
【請求項１】　データを記憶する複数のディスク装置と、
記憶すべきデータを所定の長さに分割するデータ分割手段と、
分割した所定数のデータにより構成されるパリティグループについてパリティデータを算出するパリティ算出手段と、
前記分割されたデータ及びパリティデータを前記複数の所要のディスク装置へ転送する転送手段と、
前記記憶すべきデータの格納先アドレス、サイズ、転送状況に関する情報を検出して記憶する状況記憶手段と、
復旧の際に前記状況記憶手段に記憶された情報を検索して転送が未完となっている場合には、当該格納先アドレス及びサイズに基づき転送が未完となっている領域を検出し、この領域のパリティグループについてディスク装置の記憶内容に基づきパリティデータを作成して、該当するディスク装置に書き込む復旧処理手段と、
を具備することを特徴とするディスクアレイ装置。
【請求項２】　複数のディスク装置のいずれかに障害が発生した場合には、他の正常なディスク装置に記憶されたデータに基づき障害に係るディスク装置に記憶されたデータを復元するデータ復元手段を具備することを特徴とする請求項１に記載のディスクアレイ装置。
【請求項３】　複数のディスク装置の特定の１台が、パリティデータを記憶するためのパリティディスクとなっていることを特徴とする請求項１または請求項２に記載のディスクアレイ装置。
【請求項４】　記憶すべきデータが与えられると分割し、
上記分割したデータに関するパリティグループについてパリティデータを算出し、
上記分割したデータ及びパリティデータを複数のディスク装置の所要のディスク装置に転送し、
前記記憶すべきデータの格納先アドレス、サイズ、転送状況に関する情報を検出して記憶し、
復旧の際に前記記憶された情報を検索して転送が未完となっている場合には、当該格納先アドレス及びサイズに基づき転送が未完となっている領域を検出し、この領域のパリティグループについてディスク装置の記憶内容に基づきパリティデータを作成して、該当するディスク装置に書き込むことを特徴とするデータ保守方法。
【請求項５】　複数のディスク装置のいずれかにおいて障害が発生したか否かを検出し、
障害が発生した場合には、他の正常なディスク装置に記憶されたデータに基づき障害に係るディスク装置に記憶されたデータを復元することを特徴とする請求項４に記載のデータ保守方法。
【発明の詳細な説明】
【０００１】
【発明の属する技術分野】この発明は、複数のディスク装置により構成されるディスクアレイ装置及びデータ保守方法に関するものである。
【０００２】
【従来の技術】従来、ＲＡＩＤ（Redundancy Arrays of Inexpensive Disks）法によって構成されたディスクアレイ装置が知られている。図８には、ＲＡＩＤレベル３によるディスクアレイ装置の構成例が示されている。ディスクアレイ装置２には、ＳＣＳＩ（Small Computer System Interface ）等のインタフェースバス１２を介してホスト計算機（コンピュータ）１とのデータ送受を行うためのインタフェース（ＩＦ）７が備えられている。インタフェース７からはバス１４が延びており、バス１４には、装置全体を制御する制御部２０、リード・ライトに係るデータを保持するためのデータバッファ１０、パリティ計算を行うパリティ回路９、複数の（ここでは、５台の）ディスク装置（ここでは、ＨＤＤ）との接続を行うインタフェース８－０～８－４が備えられている。
【０００３】インタフェース８－０～８－４には、ＳＣＳＩ等のインタフェースバス１３－０～１３－４を介してＨＤＤ（ハードディスクドライブ）１１－０～１１－４が接続されている。このＨＤＤ１１－０～１１－３には、データが所定のサイズに分割されて転送されて記憶され、ＨＤＤ１１－４には、上記ＨＤＤ１１－０～１１－３の対応領域のデータの排他的論理和（ＥＸ－ＯＲ）をとったパリティデータが記憶されるように構成されている。
【０００４】図８において、データＤ０～Ｄ３がそれぞれＨＤＤ１１－０～１１－３の対応する領域に格納されているものとすると、パリティデータＰは、Ｐ＝Ｄ０（ＥＸ－ＯＲ）Ｄ１（ＥＸ－ＯＲ）Ｄ２（ＥＸ－ＯＲ）Ｄ３にて得られる。そして、例えば、ＨＤＤ１１－２に障害が発生したときには、ＨＤＤ１１－２に記憶されていたデータＤ２は、Ｄ２＝Ｄ０（ＥＸ－ＯＲ）Ｄ１（ＥＸ－ＯＲ）Ｄ３（ＥＸ－ＯＲ）Ｐにて復元されて得られる。
【０００５】別の例を説明すると、図９に示されるように、ＨＤＤ１１－０～１１－３の対応する領域Ｎ、Ｎ＋１、Ｎ＋２、Ｎ＋３が５１２バイトの分割サイズを持ち、それぞれに、データ５５…、ＡＡ…、９９…、ＣＣ…（合計２０４８バイト）が記憶されるときには、パリティＰは、５５…（ＥＸ－ＯＲ）ＡＡ…（ＥＸ－ＯＲ）９９…（ＥＸ－ＯＲ）ＣＣ…＝ＡＡ…と得られ、これがＨＤＤ１１－４に書き込まれる。
【０００６】そして、上記領域Ｎ、Ｎ＋１、Ｎ＋２に対し１５３６バイトのデータが、図１０に示されるように、それぞれ、７７…、ＥＥ…、３３…と５１２バイトに分割されて書き込まれたときには、制御部２０は、領域Ｎ、Ｎ＋１、Ｎ＋２、Ｎ＋３のデータについて再度パ

リティデータを作成し直す。つまり、新たなパリティデータＰＭは、ＰＭ＝７７…（ＥＸ－ＯＲ）ＥＥ…（ＥＸ－ＯＲ）３３…（ＥＸ－ＯＲ）ＣＣ…＝６６…として得られ、これがＨＤＤ１１－４に書き込まれる。

【０００７】上記に対し、図１１に示されるように、ＨＤＤ１１－０～１１－３の対応する領域Ｎ、Ｎ＋１、Ｎ＋２、Ｎ＋３に対し５１２バイトに分割されたデータ７７…、ＥＥ…、３３…を書き込む際に、電源障害等が発生し、ＨＤＤ１１－２に対しデータ３３…の書き込みができず、正常に書き込みが終了しなかったものとする。この場合には、領域Ｎ、Ｎ＋１、Ｎ＋２には、データ７７…、ＥＥ…、３３…が書き込まれるが、領域Ｎ＋３には、元のデータＣＣ…が書き込まれた状態のまま残っており、また、ＨＤＤ１１－４の対応領域には、元のパリティデータがＡＡ…が記憶されたままとなっている。

【０００８】

【発明が解決しようとする課題】このような図１１の状態において、例えば、図１２に示すようにＨＤＤ１１－３に障害が発生し、係る状態において、ＨＤＤ１１－３の領域Ｎ＋３に対する読み出し要求があると、障害が発生しているＨＤＤ１１－３を除いたディスクの対応領域のデータを用いて復元が行われる。つまり、領域Ｎ、Ｎ＋１、Ｎ＋２のデータ７７…、ＥＥ…、３３…と、パリティデータＡＡ…を用いて、復元データ＝７７…（ＥＸ－ＯＲ）ＥＥ…（ＥＸ－ＯＲ）３３…（ＥＸ－ＯＲ）ＡＡ…＝ＡＡ…が復元され、ＨＤＤ１１－３に実際に格納されていたデータＣＣ…と異なるデータＡＡ…が復元されることになる。

【０００９】本発明は上記のようにデータの書き込み途中に生じる電源断等の異常が生じた場合に、データを適切に復元することができなくなるという問題点に鑑みなされたもので、その目的は、データの書き込み途中に生じる電源断等の異常が生じた場合にも、データを適切に復元することができるディスクアレイ装置及びデータ保守方法を提供することである。

【００１０】

【課題を解決するための手段】本発明の請求項１に係るディスクアレイ装置は、データを記憶する複数のディスク装置と、記憶すべきデータを所定の長さに分割するデータ分割手段と、分割した所定数のデータにより構成されるパリティグループについてパリティデータを算出するパリティ算出手段と、前記分割されたデータ及びパリティデータを前記複数の所要のディスク装置へ転送する転送手段と、前記記憶すべきデータの格納先アドレス、サイズ、転送状況に関する情報を検出して記憶する状況記憶手段と、復旧の際に前記状況記憶手段に記憶された情報を検索して転送が未完となっている場合には、当該格納先アドレス及びサイズに基づき転送が未完となっている領域を検出し、この領域のパリティグループについてディスク装置の記憶内容に基づきパリティデータを作成して、該当するディスク装置に書き込む復旧処理手段とを具備することを特徴とする。これにより、復旧の際に転送が未完となっている場合には、当該格納先アドレス及びサイズに基づき転送が未完となっている領域が検出され、この領域のパリティグループについてディスク装置の記憶内容に基づきパリティデータが作成されて、該当するディスク装置に書き込みが行われる。

【００１１】本発明の請求項２に係るディスクアレイ装置では、複数のディスク装置のいずれかに障害が発生した場合には、他の正常なディスク装置に記憶されたデータに基づき障害に係るディスク装置に記憶されたデータを復元するデータ復元手段を具備することを特徴とする。これにより、複数のディスク装置のいずれかに障害が発生した場合には、他の正常なディスク装置に記憶されたデータに基づき障害に係るディスク装置に記憶されたデータを復元しても適切な復元がなされる。

【００１２】本発明の請求項３に係るディスクアレイ装置では、複数のディスク装置の特定の１台が、パリティデータを記憶するためのパリティディスクとなっていることを特徴としており、これにより所定のディスク装置にパリティデータが記憶されることになる。

【００１３】本発明の請求項４に係るデータ保守装置は、記憶すべきデータが与えられると分割し、上記分割したデータに関するパリティグループについてパリティデータを算出し、上記分割したデータ及びパリティデータを複数のディスク装置の所要のディスク装置に転送し、前記記憶すべきデータの格納先アドレス、サイズ、転送状況に関する情報を検出して記憶し、復旧の際に前記記憶された情報を検索して転送が未完となっている場合には、当該格納先アドレス及びサイズに基づき転送が未完となっている領域を検出し、この領域のパリティグループについてディスク装置の記憶内容に基づきパリティデータを作成して、該当するディスク装置に書き込むことを特徴とする。これにより、複数のディスク装置のいずれかに障害が発生した場合に、他の正常なディスク装置に記憶されたデータに基づき障害に係るディスク装置に記憶されたデータを復元して適切な復元が可能となる。

【００１４】本発明の請求項５に係るデータ保守方法は、複数のディスク装置のいずれかにおいて障害が発生したか否かを検出し、障害が発生した場合には、他の正常なディスク装置に記憶されたデータに基づき障害に係るディスク装置に記憶されたデータを復元することを特徴とする。これにより、複数のディスク装置のいずれかにおいて障害が発生した場合にデータの復元がなされる。

【００１５】

【発明の実施の形態】以下添付の図面を参照して本発明の実施の形態に係るディスクアレイ装置及びデータ保守方法を説明する各図において、同一の構成要素には同一

の符号を付して重複する説明を省略する。図１には、本発明の実施の形態に係るディスクアレイ装置２Ａの構成図が示されている。この装置２Ａは、データを記憶する複数のディスク装置３１－１～３１－ｎからなるディスク部３０を有し、このディスク部３０に対し記憶すべきデータを所定の長さに分割するデータ分割手段３２と、分割した所定数のデータにより構成されるパリティグループについてパリティデータを算出するパリティ算出手段３３とを備える。更には、データ分割手段３２により分割されたデータ及びパリティ算出手段３３により作成されたパリティデータを複数の所要のディスク装置３１－１～３１－ｎへ転送する転送手段３４が設けられている。

【００１６】更に、ディスク部３０の所要のディスク装置３１－１～３１－ｎへ記憶すべきデータの格納先アドレス、サイズ、転送状況に関する情報を検出して記憶する状況記憶手段３５と、復旧の際に上記状況記憶手段３５に記憶された情報を検索して転送が未完となっている場合には、当該格納先アドレス及びサイズに基づき転送が未完となっている領域を検出し、この領域のパリティグループについてディスク装置３１－１～３１－ｎの記憶内容に基づきパリティデータを作成して、該当するディスク装置に書き込む復旧処理手段３６が備えられている。また、上記複数のディスク装置３１－１～３１－ｎのいずれかに障害が発生した場合には、他の正常なディスク装置に記憶されたデータに基づき障害に係るディスク装置に記憶されたデータを復元するデータ復元手段３７を具備する。

【００１７】上記ディスクアレイ装置２Ａの実際上の構成例を図２に示す。この構成は、図８に示したディスクアレイ装置２と同一の構成要素に同一の符号を付して示すように基本的に同一であるが、バス１４には、マイクロプロセッサ３用のＣＰＵバス１６との間のプロトコル変換を行うプロトコル変換部５が接続され、マイクロプロセッサ３がメモリ４内のプログラムやデータを用いて、上記図１に示した各手段として機能する点が特徴となっている。また、マイクロプロセッサ３は不揮発メモリ６に、ディスク部３０に対応するＨＤＤ１１－０～１１－４に記憶すべきデータの格納先アドレス、サイズ、転送状況に関する情報を検出して記憶し、状況記憶手段３５として用いる。

【００１８】図３には、上記状況記憶手段３５に係るメモリテーブルの構成例が示されている。つまり、メモリテーブルには、ホスト計算機１から送られ、ＨＤＤ１１－０～１１－４に書き込まれるデータのデータ番号（書き込み中データ１～６，・・・）、格納先アドレス（１～６，・・・）、転送サイズ（１～６，・・・）、転送の完了／未完了に係る転送状況（１～６，・・・）が記憶される。

【００１９】そして、マイクロプロセッサ３は、メモリ４に格納されている図４、図５に示されるフローチャートに対応するプログラムを用いて図１の各手段として動作するので、以下にこれを説明する。ホスト計算機１からデータの書き込み要求がなされると、マイクロプロセッサ３は、これをインタフェース７プロトコル変換部５を介して受け取り（Ｓ１）、データを受けてデータバッファ１０へ格納した後に、当該データを一定のサイズ（例えば、５１２バイト）に分割して、該当領域のパリティグループについてパリティデータを作成する（Ｓ２）。具体的には、図１０に示したように、１５３６バイトのデータが与えられたときには、それぞれ、７７…、ＥＥ…、３３…と５１２バイトに分割し、かつ、パリティグループであるＨＤＤ１１－３のデータＣＣ…を読み出し、パリティデータＰＭ＝７７…（ＥＸ－ＯＲ）ＥＥ…（ＥＸ－ＯＲ）３３…（ＥＸ－ＯＲ）ＣＣ…＝６６…を得る。

【００２０】次に、マイクロプロセッサ３は、ホスト計算機１から要求のあったデータの書き込み要求に係るデータに関し、格納先アドレス、転送サイズ、転送状況（初期値は、「未完了」）を図３のメモリテーブルに記憶する（Ｓ３）。例えば、図１０に示すように書き込みを行うときには、格納先アドレス（先頭アドレス）は「Ｎ」であり、転送サイズは「１５３６バイト」であり、転送状況の初期値は、「未完了」とされる。次に、マイクロプロセッサ３は、該当するＨＤＤに対しデータ書き込み要求を送出しデータ書き込みを開始し（Ｓ４）、当該データの書き込みが完了したかを検出する（Ｓ５）。

【００２１】上記ステップＳ５において、一連のデータ（図１０の例では、１５３６バイト）の書き込みが完了すると、図３のメモリテーブルの該当するエリアの転送状況を「未完了」から「完了」へと変更し（Ｓ６）、次の処理（ホスト計算機１からのデータの書き込み要求に係る処理）を実行する（Ｓ７）。

【００２２】上記に対し、ステップＳ５においてデータをＨＤＤ１１－０～１１－４に格納しているときに、停電等の電源断が生じたときには、動作が停止し、電源投入等による復旧がなされると、図５に示されるフローチャートのプログラムが実行される。即ち、当該図５のフローチャートのプログラムのスタート等の立ち上げ時初期処理がなされ（Ｓ１１）、次いでマイクロプロセッサ３は不揮発メモリ６内のメモリテーブルを検索して転送状況が「未完了」となっているデータがあるか否かを検出する（Ｓ１２）。「未完了」がないときには、当該ディスクアレイ装置２Ａの運用を開始する（Ｓ１３）。つまり、ホスト計算機１からの要求に応えてデータの格納または読み出しを行う。

【００２３】一方、転送状況が「未完了」となっているデータがある場合には、当該領域の格納先アドレス、転送サイズより、ＨＤＤ１１－０～１１－４において書き

込みが行われなかったパリティグループの領域を検索し、当該パリティグループのパリティデータを再作成してパリティディスクであるＨＤＤ１１－４へ当該パリティデータを書き込み（Ｓ１４）、ステップＳ１２へ戻って処理を続ける。

【００２４】例えば、図１１に示されるように、ＨＤＤ１１－０～１１－３の対応する領域Ｎ、Ｎ＋１、Ｎ＋２、Ｎ＋３に対し５１２バイトに分割されたデータ７７…、ＥＥ…、３３…を書き込む際に、電源障害等が発生し、ＨＤＤ１１－２に対しデータ３３…の書き込みができず、正常に書き込みが終了しなかったものとする。このときには、この一連のデータのデータ番号が「１」であるとき、マイクロプロセッサ３によって不揮発メモリ６のメモリテーブルには、図６（ａ）に示されるように、格納先アドレスが「Ｎ」、転送サイズが「１５３６」、転送状況が「未完了」とされる。

【００２５】上記のようなメモリテーブルの内容を図５のフローチャートにおけるステップＳ１２にて検出すると、上記格納先アドレス「Ｎ」に対応してＨＤＤ１１－０の領域Ｎに関するパリティグループのデータ７７…、ＥＥ…、９９…、ＣＣ…を読み出し、これらに基づきパリティデータＰＡ＝７７…（ＥＸ－ＯＲ）ＥＥ…（ＥＸ－ＯＲ）９９…（ＥＸ－ＯＲ）ＣＣ…を演算し、パリティデータＰＡ＝ＣＣ…を得て、これをＨＤＤ１１－４の該当領域に記憶する。この結果ＨＤＤ１１－０～１１－４には、図６（ｂ）に示されるようにデータが格納される。

【００２６】このような図６（ｂ）の状態において、例えば、ＨＤＤ１１－３に障害が発生し、係る状態において、ＨＤＤ１１－３の領域Ｎ＋３に対する読み出し要求があると、障害が発生しているＨＤＤ１１－３を除いたディスクの対応領域のデータを用いて復元が行われる。つまり、領域Ｎ、Ｎ＋１、Ｎ＋２のデータ７７…、ＥＥ…、９９…と、パリティデータＣＣ…を用いて、復元データ＝７７…（ＥＸ－ＯＲ）ＥＥ…（ＥＸ－ＯＲ）９９…（ＥＸ－ＯＲ）ＣＣ…＝ＣＣ…が復元され、ＨＤＤ１１－３に実際に格納されていたデータＣＣ…と同一のデータＣＣ…が復元されることになる。

【００２７】斯して、ＨＤＤ１１－０～１１－４に対するデータ書き込みの途中において停電等の電源断が生じて、データ書き込みが未完了となったときにも、未完了のデータ書き込みに係るパリティグループに関しパリティデータが適切なものに変更される。このため、その後にＨＤＤ１１－０～１１－４のいずれかに障害が生じても、障害に係るＨＤＤのデータを適切に復元することができることになる。なお、上記の実施の形態においては、端部に配置されたＨＤＤ１１－４にパリティデータを書き込むことにしたが、複数のＨＤＤの内のいずれかにパリティデータを書き込むように構成した実施の形態が存在する。従って、特定の１つのＨＤＤをパリティディスクとしなくとも良く、未完了のデータ書き込みに係るパリティグループに関しパリティデータを再作成して、所定の手法により決まるＨＤＤにこのパリティデータを書き込むように構成することのできる。係る場合にも、障害に係るＨＤＤのデータを適切に復元することができることになる。また、ＨＤＤはディスク装置の一例に過ぎず、他のディスク装置、例えば、光ディスク装置や光磁気ディスク装置を用いた構成を採用することも可能である。

【００２８】

【発明の効果】以上説明したように請求項１に係るディスクアレイ装置によれば、復旧の際に転送が未完となっている場合には、当該格納先アドレス及びサイズに基づき転送が未完となっている領域が検出され、この領域のパリティグループについてディスク装置の記憶内容に基づきパリティデータが作成されて、該当するディスク装置に書き込みが行われる。

【００２９】以上説明したように請求項２に係るディスクアレイ装置によれば、複数のディスク装置のいずれかに障害が発生した場合には、他の正常なディスク装置に記憶されたデータに基づき障害に係るディスク装置に記憶されたデータを復元しても適切な復元がなされる。

【００３０】以上説明したように請求項３に係るディスクアレイ装置によれば、複数のディスク装置の特定の１台が、パリティデータを記憶するためのパリティディスクとなっているので、所定のディスク装置にパリティデータを記憶して適切にデータを復元することが可能である。

【００３１】以上説明したように請求項４に係るデータ保守方法によれば、複数のディスク装置のいずれかに障害が発生した場合に、他の正常なディスク装置に記憶されたデータに基づき障害に係るディスク装置に記憶されたデータを復元して適切な復元が可能となる効果がある。

【００３２】以上説明したように請求項５に係るデータ保守方法によれば、複数のディスク装置のいずれかにおいて障害が発生したか否かを検出し、障害が発生した場合には、他の正常なディスク装置に記憶されたデータに基づき障害に係るディスク装置に記憶されたデータを復元するので、複数のディスク装置のいずれかにおいて障害が発生した場合にデータの復元がなされることになる。

【図面の簡単な説明】

【図１】本発明の実施の形態に係るディスクアレイ装置の機能ブロック図。

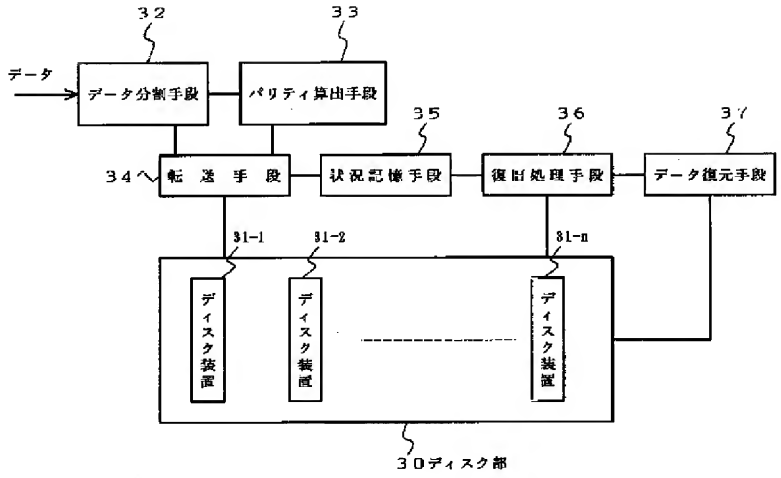
【図２】本発明の実施の形態に係るディスクアレイ装置の詳細ブロック図。

【図３】本発明の実施の形態に係るディスクアレイ装置に用いられるメモリテーブルの内容を示す図。

【図４】本発明の実施の形態に係るディスクアレイ装置

の動作を説明するためのフローチャート。
【図５】本発明の実施の形態に係るディスクアレイ装置の動作を説明するためのフローチャート。
【図６】本発明の実施の形態に係るディスクアレイ装置における電源断から復旧後の動作を説明するための図。
【図７】本発明の実施の形態に係るディスクアレイ装置におけるデータ復元動作を説明するための図。
【図８】従来のディスクアレイ装置の詳細ブロック図。
【図９】ＲＡＩＤにおけるパリティ生成を説明する図。
【図１０】ＲＡＩＤにおけるデータ書き込み及びパリティ生成を説明する図。
【図１１】従来のディスクアレイ装置おけるデータ書き込み時の電源断に際する処理動作を説明するための図。
【図１２】従来のディスクアレイ装置におけるデータ復元動作を説明するための図。
【符号の説明】
１　ホスト計算機　　２Ａ　ディスクアレイ装置
３　マイクロプロセッサ　　４　メモリ
５　プロトコル変換部　　６　不揮発性メモリ
７、８－０～８－４　インタフェース　　９　パリティ回路
１０　データバッファ　　１１－０～１１－４　ＨＤＤ
３０　ディスク部　　３１－１～３１－ｎ　ディスク装置
３２　データ分割手段　　３３　パリティ算出手段
３４　転送手段　　３５　状況記憶手段
３６　復旧処理手段　　３７　データ復元手段

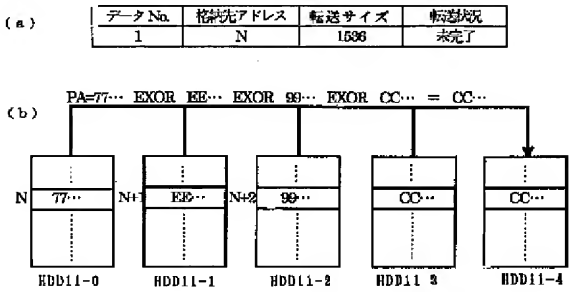
【図１】



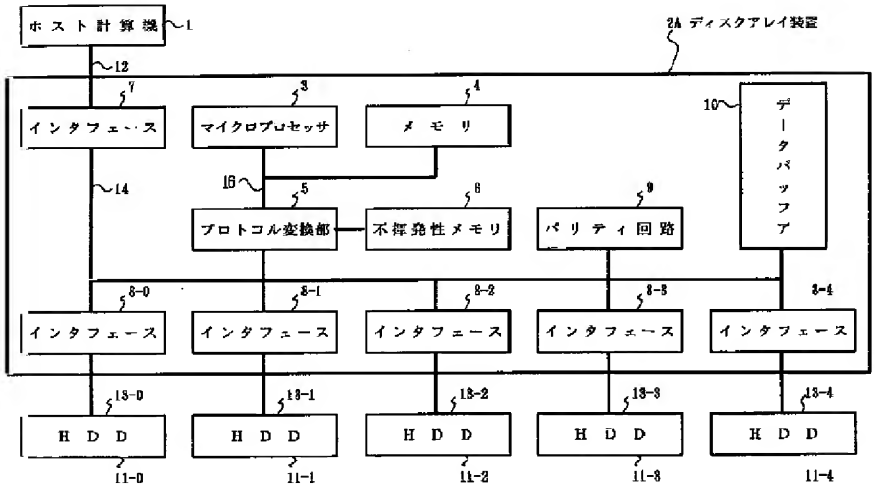

【図３】

| | | | |
|---|---|---|---|
| 書き込み中データ１ | 格納先アドレス１ | 転送サイズ１ | 転送状況１（完了／未完了） |
| 書き込み中データ２ | 格納先アドレス２ | 転送サイズ２ | 転送状況２（完了／未完了） |
| 書き込み中データ３ | 格納先アドレス３ | 転送サイズ３ | 転送状況３（完了／未完了） |
| 書き込み中データ４ | 格納先アドレス４ | 転送サイズ４ | 転送状況４（完了／未完了） |
| 書き込み中データ５ | 格納先アドレス５ | 転送サイズ５ | 転送状況５（完了／未完了） |
| 書き込み中データ６ | 格納先アドレス６ | 転送サイズ６ | 転送状況６（完了／未完了） |
| ： | ： | ： | ： |
| ： | ： | ： | ： |

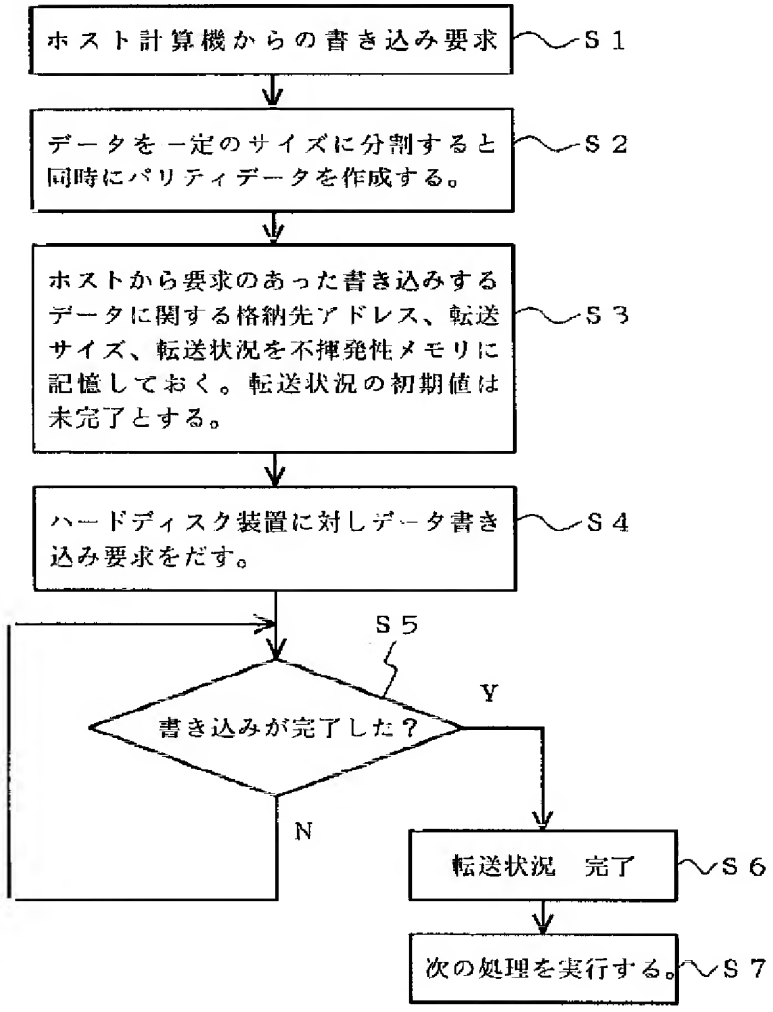
【図６】



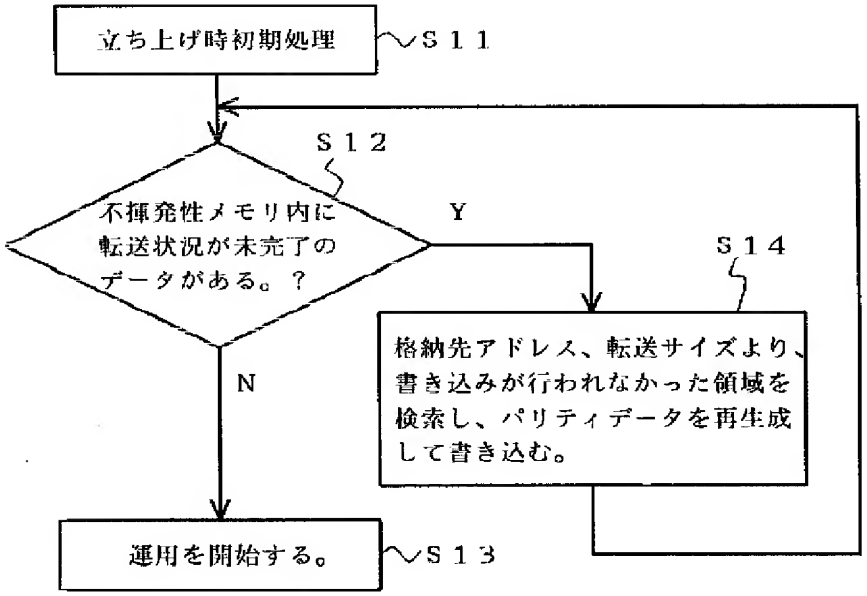
【図２】


ホスト計算機
1
2A ディスクアレイ装置
12
インタフェース
7
マイクロプロセッサ
3
メモリ
4
データバッファ
10
14
16
プロトコル変換部
5
不揮発性メモリ
6
パリティ回路
9
8-0
8-1
8-2
8-3
8-4
インタフェース
インタフェース
インタフェース
インタフェース
インタフェース
13-0
13-1
13-2
13-3
13-4
HDD
HDD
HDD
HDD
HDD
11-0
11-1
11-2
11-3
11-4

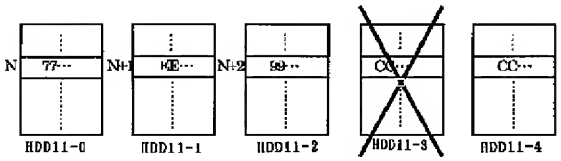
【図４】


ホスト計算機からの書き込み要求
S１
データを一定のサイズに分割すると同時にパリティデータを作成する。
S２
ホストから要求のあった書き込みするデータに関する格納先アドレス、転送サイズ、転送状況を不揮発性メモリに記憶しておく。転送状況の初期値は未完了とする。
S３
ハードディスク装置に対しデータ書き込み要求をだす。
S４
S５
書き込みが完了した？
Y
N
転送状況　完了
S６
次の処理を実行する。
S７

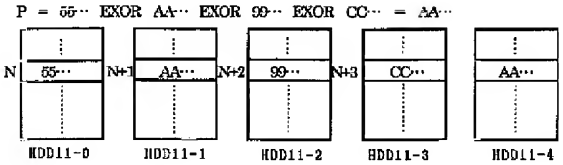
【図５】


立ち上げ時初期処理
S１１
S１２
不揮発性メモリ内に転送状況が未完了のデータがある。？
Y
N
S１４
格納先アドレス、転送サイズより、書き込みが行われなかった領域を検索し、パリティデータを再生成して書き込む。
運用を開始する。
S１３

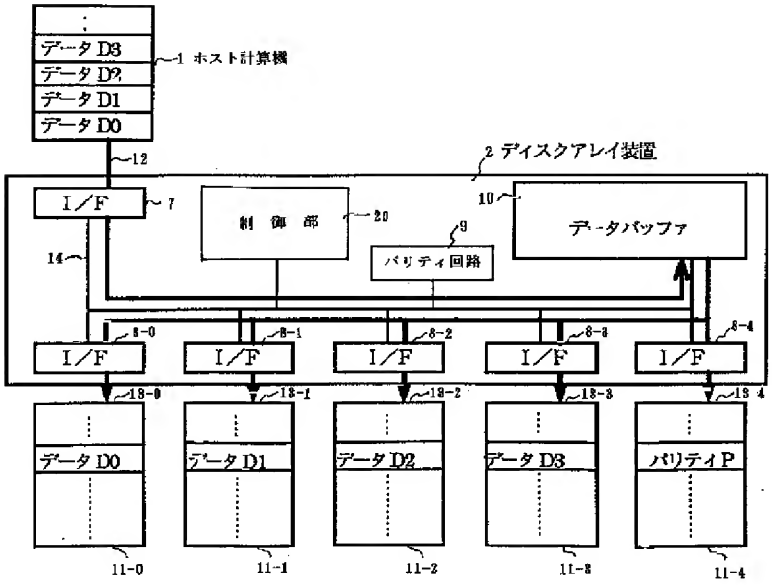
【図７】


N
77…
N+1
6E…
N+2
99…
CC…
CC…
HDD11-0
HDD11-1
HDD11-2
HDD11-3
HDD11-4

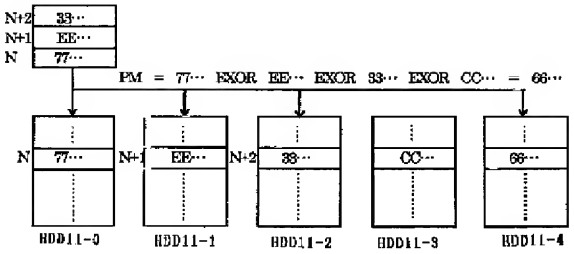
【図９】


P = 55… EXOR AA… EXOR 99… EXOR CC… = AA…
N
55…
N+1
AA…
N+2
99…
N+3
CC…
AA…
HDD11-0
HDD11-1
HDD11-2
HDD11-3
HDD11-4

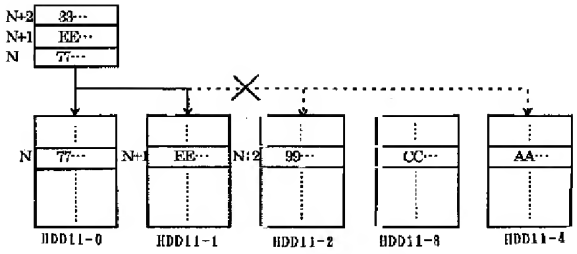
【図８】


データ D3
データ D2
データ D1
データ D0
1 ホスト計算機
12
2 ディスクアレイ装置
I／F
7
制 御 部
20
パリティ回路
9
データバッファ
10
14
8-0
8-1
8-2
8-3
8-4
I／F
I／F
I／F
I／F
I／F
18-0
18-1
18-2
18-3
18-4
データ D0
データ D1
データ D2
データ D3
パリティ P
11-0
11-1
11-2
11-3
11-4

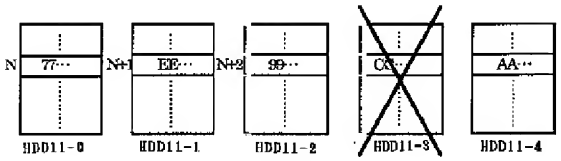
【図１０】


N+2
33…
N+1
EE…
N
77…
PM = 77… EXOR EE… EXOR 33… EXOR CC… = 66…
N
77…
N+1
EE…
N+2
33…
CC…
66…
HDD11-0
HDD11-1
HDD11-2
HDD11-3
HDD11-4

【図１１】


N+2
33…
N+1
EE…
N
77…
N
77…
N+1
EE…
N+2
99…
CC…
AA…
HDD11-0
HDD11-1
HDD11-2
HDD11-3
HDD11-4

【図１２】


N
77…
N+1
EE…
N+2
99…
CC…
AA…
HDD11-0
HDD11-1
HDD11-2
HDD11-3
HDD11-4